Discrete Mathematics 75 (1989) 255–278
North-Holland

PSEUDO-RANDOM HYPERGRAPHS

Julie HAVILAND and Andrew THOMASON
Department of Pure Mathematics and Mathematical Statistics, 16, Mill Lane, Cambridge CB2 1SB, England

1. Introduction

The study of random graphs has proved very successful in showing the existence of graphs which are extremal with respect to certain properties (see Bollobás [1] for a detailed exposition). Typical of the problems to which they have been applied are subcontractions [11], Zarankiewicz's problem [10] and Ramsey's theorem [6]. Random graphs also offer us examples of graphs with particular properties, giving us expanders [4], graphs of small diameter [3] and parallel sorting algorithms [2]. In most cases the difficulty remains of constructing explicit extremal graphs, or of checking whether a given randomly-generated graph is extremal. As an initial approach to this problem, a simple criterion was proposed in [12], whereby any graph satisfying it might be regarded as a *pseudo-random graph*; that is, it would possess certain desirable properties of random graphs. The criterion was stated in the following terms: a graph G is said to be (p, α)*-jumbled* if p and α are real numbers satisfying $0<p<1\leqslant\alpha$, and if every induced subgraph H of G satisfies

$$\left|e(H)-p\binom{|H|}{2}\right|\leqslant\alpha\,|H|,$$

where e(H) is the number of edges in H. Other possible definitions of pseudo-random graphs are offered by Chung, Graham and Wilson in [5]; the definitions turn out to be roughly compatible, but for our purposes the definition given will prove the most satisfactory.

A (p, α)-jumbled graph could be regarded as behaving rather like a random graph of edge-probability p, the parameter α determining the closeness of this resemblance. In fact a modification of a theorem of Erdős and Spencer [8] shows that α must be at least of order $|G|^{\frac{1}{2}}$, and subject to this constraint it is not hard to verify that almost all random graphs with edge probability p are (p, α)-jumbled.

In [12] and [13] it was shown that (p, α)-jumbled graphs possess some of the desirable properties of random graphs, at least for reasonably large values of p ($p\gg n^{-\frac{1}{2}}$, and usually $p\gg n^{-\frac{1}{3}}$). Moreover two sufficient conditions were found for a graph to be (p, α)-jumbled. One, stated in Proposition B below, is a 'global' condition on all induced subgraphs of a fixed size (rather than all sizes as in the definition). The other, stated in Proposition A below, is an easily checked 'local' condition on the degrees of the vertices and of vertex pairs. This latter theorem is a means whereby a specific graph can be shown to be (p, α)-jumbled.

The remarks about the value of random graphs equally apply to random hypergraphs, and our aim in this paper is to try to extend the work on (p, α)-jumbled graphs to r-uniform hypergraphs. A specific motive for doing this is to answer a question of Erdős and Sós, posed in [7], concerning the number of K_4^3s (complete 3-uniform hypergraphs on 4 vertices) in 3-uniform hypergraphs. A solution to their question would follow immediately if the obvious generalisation to 3-hypergraphs of a theorem in [12], giving the number of cliques in a jumbled graph, were true. It turns out, rather surprisingly, that this particular generalisation is false. Nevertheless, we are able to prove characterisation theorems analogous to Propositions A and B, and to establish some of the basic properties of jumbled hypergraphs.

To begin, we shall define a (p, α)-jumbled hypergraph in the obvious way.

Definition. An r-uniform $(r \geq 3)$ hypergraph G is said to be (p, α)-*jumbled* if p, α are real numbers satisfying $0 < p < 1 \leq \alpha$, and if every induced (r-uniform) subgraph H of G satisfies

$$\left| e(H) - p\binom{|H|}{r} \right| \leq \alpha\, |H|,$$

where $e(H)$ is the number of edges in H.

Adding extra conditions in this definition might permit stronger theorems to be proved. However we shall not do this, since it would defeat the object of the exercise, which is to see what consequences follow from just this simple definition.

Notation

We shall employ throughout the following notation. If x is a nonnegative integer, then $B(x)$ will denote any real number y of absolute value at most x. Hence $y = B(x)$ means $|y| \leq x$, and $0 \leq z \leq x$ implies $B(z) = B(x)$. In this sense the notation behaves like Landau's $O(x)$ notation. Therefore we may rewrite the definition of a (p, α)-jumbled r-uniform hypergraph G as

$$e(H) = p\binom{|H|}{r} + B(\alpha\, |H|)$$

for all induced $H \subset G$. Further, all hypergraphs will be r-uniform for some r, and we shall often refer to a hypergraph as *jumbled* if it is (p, α)-jumbled for some p and α whose actual values are not of specific interest. In fact the extension of

Erdős and Spencer's theorem to hypergraphs shows α must be of order at least $|G|^{(r-1)/2}$. (A full proof of this fact, together with the less illuminating details of other later proofs, is presented in [9]). We use $X^{(t)}$ to denote $\{Y \subset X; |Y| = t\}$, and if x is a real number, $(x)_t$ denotes the falling factorial $x(x-1)\cdots(x-t+1)$. We shall say that $\sigma \in V(G)^{(r-1)}$ and $x \in V(G)$ are neighbours if $\sigma \cup \{x\} \in E(G)$. If H is an induced subgraph of G, we write $d_H(x)$, the degree of x in H, for $|\{\tau \in V(H)^{(r-1)}; \tau \cup \{x\} \in E(G)\}|$ and $d_H(\sigma)$, the degree of σ in H, for $|\{y \in V(H); \sigma \cup \{y\} \in E(G)\}|$. If $H = G$ the subscript may be omitted. Finally, if $S \subset V(G)$ and $T \subset V(G) \backslash S$, then the set of edges of G containing i vertices of S and $r-i$ of T will be denoted by $E_i(S, T)$, and we write $e_i(S, T)$ for $|E_i(S, T)|$. If H and F are induced subgraphs of G, we may write $e_i(H, F)$ instead of $e_i(V(H), V(F))$.

Examples of pseudo-random hypergraphs

We now give just a few examples of pseudo-random hypergraphs, some of which generalize examples of [12]; another appears later in the paper. Verification of the examples can be found in [9].

(1) Almost all r-uniform hypergraphs G, having edges chosen independently with probability p, are $(p, O(|G|^{(r-1)/2}))$-jumbled. This is a straightforward exercise in random hypergraph theory. Alternatively, Theorem 1 below can be used to show G is $(p, O(|G|^{r-\frac{5}{4}}))$-jumbled.

(2) Let q be a prime and let $\mathbb{F}_q$ be the field of order q. Consider the hypergraph G where $V(G) = \mathbb{F}_q$ and $\{x_1, \ldots, x_r\} \in E(G)$ if and only if $x_1 + \cdots + x_r$ is a square $(\operatorname{mod} q)$. Elementary theory of characters over finite fields shows that for this graph, each vertex appears in a number of edges in the range $\frac{1}{2}\binom{q-1}{r-1} \pm \dfrac{r(q-1)_{r-2}}{2(r-1)!}$. Moreover, for each pair $x, y \in V(G)$, their number of common neighbours lies in the range $\frac{1}{4}\binom{q-2}{r-1} \pm \dfrac{3(r+1)(q-2)_{r-2}}{4(r-1)!}$. It will follow from Theorem 1 below that G is $(\frac{1}{2}, 2\,|G|^{r-\frac{3}{2}})$-jumbled.

(3) Let the vertices of G be the q^{2k} vectors in a vector space V of dimension $2k$ over $\mathbb{F}_q$, and let f be a non-degenerate quadratic form on V. Let $\{x_1, \ldots, x_r\} \in E(G)$ if and only if $f(x_1 + \cdots + x_r) = 0$. Again Theorem 1 below can be applied to show G is $(1/q, 2\,|G|^{r-\frac{5}{4}})$-jumbled.

2. Conditions implying a hypergraph is jumbled

In [12], the two propositions stated below provided local and global tests respectively for determining whether specific graphs were (p, α)-jumbled.

Proposition A ([12]). *Let n be an integer, and let $0<p<1$ and $\mu \geq 0$ be real numbers. If G is a graph of order n with minimum degree pn in which no two vertices have more than $(p^2+\mu)n$ common neighbours, then G is (p, α)-jumbled, where $\alpha^2 = n(p+\mu n)$.*

Proposition B ([12]). *Let p, η, α, n, ω be positive real numbers with $0<p, \eta<1$ such that ηn is an integer with $2 \leq \eta n \leq n-2$. Let G be a graph of order n in which for every induced subgraph H of order ηn, $|e(H) - p\binom{\eta n}{2}| \leq \eta n \alpha$ holds. Then G contains a subgraph G^* of order at least $(1-880\eta^{-1}(1-\eta)^{-2}\omega^{-1})n$ which is $(p, \omega\alpha)$-jumbled.*

Our analogue of Proposition A, a local test for checking whether a given hypergraph is (p, α)-jumbled, is Theorem 1 below.

Theorem 1. *Let n and r be integers, $r \leq 2r-3 \leq n$, and let $\delta \geq 1$, μ and $p<1$ be positive real numbers. If G is an r-uniform hypergraph of order n with every pair of vertices having at most $(p^2+\mu)\binom{n-2}{r-1}$ common neighbours, and with the number of neighbours of every $(r-1)$-set lying in the range $\{p(n-r+1), p(n-r+1)+\delta\}$, then G is (p, α)-jumbled, where*

$$\alpha^2 = \frac{1}{r!} n^{2r-3}[p(1-p) + \mu(n-r) - 10\delta(p+\delta/n)].$$

Proof. Let H be an induced subgraph of G of order k. We shall assume $r \leq k \leq n-r$, otherwise the result is easily checked. Let $d = \binom{k}{r-1}^{-1} \sum d_H(\sigma)$, the sum being over all $\sigma \in V(H)^{(r-1)}$. Thus

$$e(H) = \frac{1}{r}\binom{k}{r-1} d. \tag{1}$$

For each i, $1 \leq i \leq r$, we abbreviate $e_i(H, G-H)$ to e_i; note $e_r = e(H)$. Denote by X_j the set $\{\sigma \in V(G)^{(r-1)}; |\sigma \cap V(H)| = r-1-j\}$. By summing $d_G(\sigma)$ for $\sigma \in X_j$, and using the bounds on $d_G(\sigma)$ given by the conditions of the theorem, we obtain the inequalities

$$\begin{aligned}\binom{k}{r-1-j}\binom{n-k}{j} p(n-r+1) &\leq (r-j)e_{r-j} + (j+1)e_{r-1-j} \\ &\leq \binom{k}{r-1-j}\binom{n-k}{j}\{p(n-r+1)+\delta\},\end{aligned}$$

for each $j<r$. From these inequalites we recover lower and upper bounds $b_i^{\min}$ and $b_i^{\max}$ for the quantity ie_i. Substituting (1) into the inequality with $j=0$ gives

$$\begin{aligned}b_{r-1}^{\min} = (r-1)\binom{k}{r-1}\{p(n-r+1)-d\} &\leq (r-1)e_{r-1} \\ &\leq (r-1)\binom{k}{r-1}\{p(n-r+1)-d+\delta\} = b_{r-1}^{\max}.\end{aligned}$$

These bounds can be substituted into the inequality with $j = 1$, the results from that being substituted into the inequality with $j = 2$, and so on. In general, write

$$S(j) = \binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} k \sum_{i=1}^{j} (-1)^i \binom{n-r+1}{k-r+i}.$$

Also, setting $A_j = \{i; 1 \leqslant i \leqslant j,\ i \not\equiv j \pmod 2\}$ and $B_j = \{i; 1 \leqslant i \leqslant j,\ i \equiv j \pmod 2\}$, write

$$T(j) = \frac{k}{n-r+1}\binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i \in A_j} \binom{n-r+1}{k-r+i},$$

and

$$U(j) = \frac{k}{n-r+1}\binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i \in B_j} \binom{n-r+1}{k-r+i};$$

then it can be verified by induction on j that, for $j \geqslant 0$,

$$b_{r-j}^{\min} = (-1)^j \binom{r-1}{j}\left[pS(j) + \binom{k}{r-1}d - (-1)^j \delta T(j)\right],$$

and

$$b_{r-j}^{\max} = (-1)^j \binom{r-1}{j}\left[pS(j) + \binom{k}{r-1}d + (-1)^j \delta U(j)\right].$$

In fact we shall not use $b_j^{\max}$ in the sequel, but it was required to obtain $b_{j-1}^{\min}$. We shall assume $b_j^{\min} \geqslant 0$. This will be the case for values of δ of practical interest; nevertheless the theorem holds despite this constraint, as shown in [9].

Now, since every pair of vertices in G has at most $(p^2 + \mu)\binom{n-2}{r-1}$ common neighbours, then summing over all pairs of vertices in H, we have

$$\binom{k}{2}(p^2 + \mu)\binom{n-2}{r-1} \geqslant \sum_{\sigma \in V(G)^{(r-1)}} \binom{d_H(\sigma)}{2} = \sum_{j=0}^{r-1} \sum_{\sigma \in X_j} \binom{d_H(\sigma)}{2}.$$

Note that $\sum_{\sigma \in X_j} d_H(\sigma) = (r-j)e_{r-j} \geqslant b_{r-j}^{\min} \geqslant 0$, so

$$\binom{k}{2}(p^2 + \mu)\binom{n-2}{r-1} \geqslant \sum_{j=0}^{r-1} |X_j| \binom{b_{r-j}^{\min} |X_j|^{-1}}{2}.$$

Multiplying through by two and rearranging as a quadratic in d, we get

$$\begin{aligned} d^2\binom{k}{r-1}^2 \sum_{j=0}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} + 2\,dp\binom{k}{r-1} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 S(j)\,|X_j|^{-1} \\ -2\,d\delta\binom{k}{r-1} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j}^2 T(j)\,|X_j|^{-1} \\ \leqslant p^2\left[k(k-1)\binom{n-2}{r-1} - \sum_{j=0}^{r-1} \binom{r-1}{j}^2 S(j)^2\,|X_j|^{-1}\right] + p \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} S(j) \\ + \mu k(k-1)\binom{n-2}{r-1} - \delta \sum_{j=0}^{r-1} \binom{r-1}{j}\left[1 - (-1)^j \binom{r-1}{j} 2pS(j)\,|X_j|^{-1}\right] T(j) \\ - \delta^2 \sum_{j=0}^{r-1} \binom{r-1}{j}^2 T(j)^2\,|X_j|^{-1}. \end{aligned}$$

It is demonstrated in the Appendix that the following identities hold:

$$\binom{k}{r-1}\sum_{j=0}^{r-1}\binom{r-1}{j}^2|X_j|^{-1}=-\frac{1}{k-r+1}\sum_{j=0}^{r-1}\binom{r-1}{j}^2S(j)\,|X_j|^{-1}$$
$$=\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1},$$

$$\sum_{j=0}^{r-1}(-1)^j\binom{r-1}{j}S(j)=k\binom{n-1}{r-1},\quad\text{and}$$

$$\sum_{j=0}^{r-1}\binom{r-1}{j}^2S(j)^2\,|X_j|^{-1}-k(k-1)\binom{n-2}{r-1}$$
$$=k\binom{n-1}{r-1}+(k-r+1)^2\binom{k}{r-1}\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1}.$$

Using these identities and writing

$$C=\binom{n-k}{r-1}\binom{n-r+2}{r-1}^{-1}\sum_{j=0}^{r-1}(-1)^j\binom{r-1}{j}^2T(j)\,|X_j|^{-1},$$

we complete the square for d and obtain

$$\binom{k}{r-1}\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1}[d-p(k-r+1)+C\delta]^2$$
$$\leqslant p(1-p)k\binom{n-1}{r-1}+\mu k(k-1)\binom{n-2}{r-1}$$
$$-\delta\sum_{j=0}^{r-1}\binom{r-1}{j}\left[1-(-1)^j2p\,|X_j|^{-1}\binom{r-1}{j}\left(S(j)-(k-r+1)\binom{k}{r-1}\right)\right]T(j)$$
$$-\delta^2\sum_{j=0}^{r-1}\binom{r-1}{j}^2T(j)^2\,|X_j|^{-1}+\binom{k}{r-1}\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1}\delta^2C^2.$$

This inequality has the form

$$(d-p(k-r+1)+C\delta)^2\leqslant D.$$

For real numbers a, b and c with $b, c>0$, it can be verified that $(a+b)^2\leqslant c^2$ implies $a^2\leqslant(c+b)^2\leqslant 2(c^2+b^2)$, from which it can be shown that $(a+b)^2\leqslant c^2$ implies $a^2\leqslant 2(c^2+b^2)$ for all a, b, c. Thus we have

$$(d-p(k-r+1))^2\leqslant 2(D+(C\delta)^2),$$

yielding

$$\left|e(H)-p\binom{k}{r}\right|^2=\frac{1}{r^2}\binom{k}{r-1}^2(d-p(k-r+1))^2$$
$$\leqslant\frac{2}{r^2(r-1)^2}\binom{k-1}{r-2}^2(D+(C\delta)^2)\,|H|^2$$
$$\leqslant\frac{2\,|H|^2}{r^2(r-1)}\binom{k-1}{r-2}\binom{n-k}{r-1}\binom{n-r+2}{r-1}^{-1}\left\{\binom{n-1}{r-1}\left[p(1-p)+\frac{\mu(k-1)(n-r)}{(n-1)}\right]\right.$$
$$+\frac{\delta}{k}\sum_{j=0}^{r-1}\binom{r-1}{j}|T(j)|+\frac{2\delta p}{k}\sum_{j=0}^{r-1}\binom{r-1}{j}^2|T(j)|\,|X_j|^{-1}\left(|S(j)|+r\binom{k}{r}\right)$$
$$\left.+\frac{\delta^2}{k}\sum_{j=0}^{r-1}\binom{r-1}{j}^2|T(j)|^2\,|X_j|^{-1}+\frac{2\delta^2}{k}\binom{k}{r-1}\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1}C^2\right\},$$

which we denote in the obvious way as

$$\left|e(H)-p\binom{k}{r}\right|^2\leqslant|H|^2\,M\{E_1+E_2+E_3+E_4+E_5\}. \tag{2}$$

We now give bounds for $S(j)$, $T(j)$ and $|X_j|$, and so for ME_i, $1\leqslant i\leqslant 5$. It is easily checked that

$$|S(j)|\leqslant\frac{k(n-r+1)}{(r-1)!}\sum_{i=1}^{j}(k-1)_{r-i-1}(n-k)_{i-1}\leqslant\frac{k(n-1)^{r-1}}{(r-2)!},$$
$$|T(j)|\leqslant\frac{k}{(r-1)!}\sum_{i\in A_i}(k-1)_{r-i-1}(n-k)_{i-1}\leqslant\frac{k(n-1)^{r-2}}{2(r-2)!},\quad\text{and}$$
$$\binom{n-k}{r-1}\binom{k-1}{r-2}\binom{r-1}{j}|X_j|^{-1}=\frac{(n-k-j)_{r-1-j}(k-r+j+1)_j}{k(r-2)!}\leqslant\frac{(n-k)^{r-1-j}k^{j-1}}{(r-2)!}.$$

Therefore,

$$ME_1\leqslant\frac{n^{2r-3}}{r!}[p(1-p)+\mu(n-r)],$$
$$ME_2\leqslant\frac{2}{r^2(r-1)}\binom{k-1}{r-2}\binom{n-k}{r-1}\binom{n-r+2}{r-1}^{-1}\frac{\delta}{k}\frac{k(n-1)^{r-2}}{2(r-2)!}2^{r-1}$$
$$\leqslant\frac{\delta}{n}\frac{n^{2r-3}}{r!}\frac{2^{r-1}}{r(r-2)!},$$
$$ME_3\leqslant\frac{8\delta p}{kr^2(r-1)}\binom{n-r+2}{r-1}^{-1}\frac{k(n-1)^{r-2}}{2(r-2)!}\frac{k(n-1)^{r-1}}{(r-2)!}\sum_{j=0}^{r-1}\binom{r-1}{j}\frac{(n-k)^{r-1-j}k^{j-1}}{(r-2)!}$$
$$=\frac{4\delta p}{r^2(r-1)}\binom{n-r+2}{r-1}^{-1}\frac{(n-1)^{2r-3}}{(r-2)!^3}n^{r-1}$$
$$\leqslant\frac{4\delta pn^{2r-3}}{r^2(r-2)!^2}\left[1+\frac{r-2}{n-r+1}\right]^{r-2}$$
$$\leqslant\delta p\frac{n^{2r-3}}{r!}\frac{2^r(r-1)}{r(r-2)!},$$

$$ME_4 \leq \frac{2}{r^2(r-1)} \binom{n-r+2}{r-1}^{-1} \frac{\delta^2}{k} \frac{k^2(n-1)^{2r-4}}{4(r-2)!^2} \sum_{j=0}^{r-1} \binom{r-1}{j} \frac{(n-k)^{r-1-j}k^{j-1}}{(r-2)!}$$
$$\leq \frac{\delta^2(n-1)^{2r-4}n^{r-1}}{2r^2(r-1)(r-2)!^3} \binom{n-r+2}{r-1}^{-1}$$
$$\leq \frac{\delta^2}{n} \frac{n^{2r-3}}{r!} \frac{2^{r-3}(r-1)}{r(r-2)!}, \quad \text{and}$$
$$ME_5 \leq \frac{4}{r^2(r-1)^2} \delta^2 \binom{k-1}{r-2}^2 C^2$$
$$\leq \frac{4\delta^2}{r^2(r-1)^2} \binom{n-r+2}{r-1}^{-2} \frac{k^2(n-1)^{2r-4}}{4(r-2)!^2} \left[\sum_{j=0}^{r-1} \binom{r-1}{j} \frac{(n-k)^{r-1-j}k^{j-1}}{(r-2)!} \right]^2$$
$$\leq \frac{\delta^2(n-1)^{2r-4}n^{2r-2}}{r^2(r-1)^2(r-2)!^4} \binom{n-r+2}{r-1}^{-2}$$
$$\leq \frac{\delta^2}{n} \frac{n^{2r-3}}{r!} \frac{2^{2r-4}(r-1)}{r(r-2)!}.$$

The proof is completed by substituting these bounds for ME_i into (2), and noting that both

$$\frac{2^r(r-1)}{r(r-2)!} \quad \text{and} \quad \frac{2^{r-1}+(2^{r-3}+2^{r-4})(r-1)}{r(r-2)!}$$

are less than 10. □

The obvious difference between Proposition A and Theorem 1 is that in the latter we require an upper bound on the $(r-1)$-set degrees for the proof to work. It may be that the dependence of α on δ could be removed by a more careful argument in the early stages of the proof. More important, however, is that Theorem 1 never enables us to show that a hypergraph of order n is $(p, O(n^{(r-1)/2}))$-jumbled, the theoretical minimum; the best it affords is $(p, O(n^{r-\frac{3}{2}}))$-jumbled. this is a marked difference from the case $r=2$.

The remainder of this section provides a test for determining if a hypergraph is jumbled if we know the number of edges in subgraphs of a large fixed order. Hence Theorem 5 generalises Proposition B at the start of this section. To begin with, we prove a technical lemma which will be used heavily later.

Lemma 2. *Let $m \in \mathbb{N}$ and $z \in \mathbb{R}$ be positive, and suppose $x_0, \ldots, x_r$ satisfy*

$$\sum_{j=0}^{r} (im)_j x_j = B(z) \quad \textit{for } i = 0, \ldots, r.$$

Then

$$x_j = \left(\frac{2e^2}{jm}\right)^j B(2^r z) \quad \textit{for } j = 0, \ldots, r.$$

Proof. Let us fix some value of j, $0 \leqslant j \leqslant r$, and solve for x_j. We need to find numbers y_i, $0 \leqslant i \leqslant r$, such that

$$\sum_{i=0}^{r} (im)_j y_i = 1, \qquad \sum_{i=0}^{r} (im)_k y_i = 0 \quad \text{for } k \neq i, \tag{1}$$

in which case we get

$$x_j = \sum_{i=0}^{r} y_i B(z) = B\Big(z \sum_{i=0}^{r} |y_i|\Big). \tag{2}$$

Solving the Eqs. (1) by Cramer's rule, we see $y_i = \Delta_i / \Delta$; here

$$\Delta = \begin{vmatrix} (0m)_0 & \cdots & (rm)_0 \\ \vdots & \cdot & \vdots \\ (0m)_r & \cdots & (rm)_r \end{vmatrix} = D(0m, \ldots, rm),$$

where

$$D(a_0, \ldots, a_r) = \begin{vmatrix} (a_0)_0 & \cdots & (a_r)_0 \\ \vdots & \cdot & \vdots \\ (a_0)_r & \cdots & (a_r)_r \end{vmatrix} = \begin{vmatrix} a_0^0 & \cdots & a_r^0 \\ \vdots & \cdot & \vdots \\ a_0^r & \cdots & a_r^r \end{vmatrix} = \prod_{0 \leqslant l < l' \leqslant r} (a_l - a_{l'})$$

and

$$\Delta_i = \begin{vmatrix} (0m_0) & \cdots & ((i-1)m)_0 & 0 & ((i+1)m)_0 & \cdots & (rm)_0 \\ \vdots & \cdot & \vdots & \vdots & \vdots & \cdot & \vdots \\ (0m)_{j-1} & \cdots & ((i-1)m)_{j-1} & 0 & ((i+1)m)_{j-1} & \cdots & (rm)_{j-1} \\ (0m)_j & \cdots & ((i-1)m)_j & 1 & ((i+1)m)_j & \cdots & (rm)_j \\ (0m)_{j+1} & \cdots & ((i-1)m)_{j+1} & 0 & ((i+1)m)_{j+1} & \cdots & (rm)_{j+1} \\ \vdots & \cdot & \vdots & \vdots & \vdots & \cdot & \vdots \\ (0m)_r & \cdots & ((i-1)m)_r & 0 & ((i+1)m)_r & \cdots & (rm)_r \end{vmatrix}.$$

Now

$$\Delta_i = \sum_{l=0}^{j} \lambda_l D(0m, \ldots, (i-1)m, l, (i+1)m, \ldots, rm)$$

provided the λ_l satisfy

$$\sum_{l=0}^{j} \lambda_l (l)_j = 1, \qquad \sum_{l=0}^{j} \lambda_l (l)_k = 0, \quad 0 \leqslant k < j. \tag{3}$$

This means $\lambda_l = (-1)^{j+l}\binom{j}{l}/j!$, so

$$\begin{aligned} y_i &= \frac{1}{j!} \sum_{l=0}^{j} (-1)^{j+l} \binom{j}{l} \frac{D(0m, \ldots, (i-1)m, l, (i+1)m, \ldots, rm)}{D(0m, \ldots, (i-1)m, im, (i+1)m, \ldots, rm)} \\ &= \frac{1}{j!} \sum_{l=0}^{j} (-1)^{j+l} \binom{j}{l} \Bigg[(-1)^{r-i} \prod_{\substack{0 \leqslant k \leqslant r \\ (k \neq i)}} (km - l) \Bigg] \Bigg[(-1)^{r-i} \prod_{\substack{0 \leqslant k \leqslant r \\ (k \neq i)}} (km - im)^{-1} \Bigg] \\ &= \frac{1}{j!} \frac{(-1)^i m^{-r}}{i!\,(r-i)!} \sum_{l=0}^{j} (-1)^{j+l} \binom{j}{l} \prod_{\substack{0 \leqslant k \leqslant r \\ (k \neq i)}} (km - l). \end{aligned}$$

Next, define S_t to be the sum, over all products of t distinct elements from the set $\{km; 0 \leqslant k \leqslant r, k \neq i\}$. Then, expanding the product in the expression for y_i gives

$$y_i = \frac{(-1)^i m^{-r}}{i!\,(r-i)!} \sum_{u=0}^{r} S_{r-u} \sum_{l=0}^{j} \frac{(-1)^{j+l}}{j!} \binom{j}{l} l^u .$$

Observe it follows at once from (3) that the final sum vanishes if $0 \leqslant u < j$. Since the sum S_t involves $\binom{r}{t}$ terms, each at most $m^t r!/(r-t)!$, we may bound $|y_i|$ by

$$\begin{aligned}
|y_i| &\leqslant \frac{m^{-r}}{i!\,(r-i)!} \sum_{u=j}^{r} \frac{m^{r-u} r!}{u!} \sum_{l=0}^{j} \frac{l^u}{j!} \binom{j}{l} \\
&\leqslant m^{-j} \binom{r}{i} \sum_{u=j}^{r} \frac{j^u 2^j}{u!\,j!} \\
&\leqslant m^{-j} \frac{2^j}{j!} \binom{r}{i} \sum_{u=0}^{\infty} \frac{j^u}{u!} \\
&= m^{-j} \frac{(2e)^j}{j!} \binom{r}{i} \leqslant \left(\frac{2e^2}{jm}\right)^j \binom{r}{i}.
\end{aligned}$$

Finally, returning to Eq. (2),

$$x_j = B\left(z \sum_{i=0}^{r} |y_i|\right) = \left(\frac{2e^2}{jm}\right)^j B(2^r z). \quad \square$$

We are now ready to begin the work leading to the proof of Theorem 5.

Lemma 3. *Let $r \geqslant 3$ be a positive integer, and let C, n, p and η be positive real numbers with $p, \eta < 1 \leqslant C$, such that ηn is an integer with $2r \leqslant \eta n \leqslant n - 2r$. Let G be an r-uniform hypergraph of order n in which for every induced subgraph H of order ηn, $|e(H) - p\binom{\eta n}{r}| \leqslant C$ holds. Then $|e(H) - p\binom{k}{r}| \leqslant e^{6r} C \eta^{-r} (1-\eta)^{-r}$ for each induced subgraph H of order k.*

Proof. Let H be a subgraph of order $k \geqslant \eta n$. If we count the number of edges in each of the $\binom{k}{l}$ subgraphs L of H of order $l = \eta n$, we get

$$e(H) = \binom{k-r}{l-r}^{-1} \sum_{L \subset H} e(L) = \binom{k-r}{l-r}^{-1} \sum_{L \subset H} \left[p\binom{l}{r} + B(C) \right] = p\binom{k}{r} + \frac{(k)_r}{(l)_r} B(C).$$

Observe that

$$\begin{aligned}
\frac{(k)_r}{(l)_r} = \frac{k^r}{l^r} \prod_{i=1}^{r-1} \frac{\left(1-\frac{i}{k}\right)}{\left(1-\frac{i}{l}\right)} &\leqslant \frac{k^r}{l^r}\left(1-\frac{r}{l}\right)^{-r} \leqslant \frac{k^r}{l^r}\left(1-\frac{2r}{l}\right)^{r} \quad \text{for } l = \eta n \geqslant 2r \\
&\leqslant \eta^{-r} e^{2r^2/l} \leqslant \eta^{-r} e^r .
\end{aligned}$$

Hence

$$e(H)=p\binom{k}{r}+B(e^rC\eta^{-r})=p\binom{k}{r}+B\left(\frac{e^{6r}}{\eta^r(1-\eta)^r}C\right).$$

The lemma holds easily for $k\leqslant 2r$; now suppose H is a subgraph of order $2r\leqslant k\leqslant\min\{(1-\eta)n,\eta n\}$. Let F be a subgraph of $G-H$ of order ηn, and let L be a subgraph of H of order l, where $1\leqslant l\leqslant k$. Then by the above,

$$e(L\cup F)=p\binom{l+\eta n}{r}+B(e^rC\eta^{-r}).$$

Summing over all $\binom{k}{l}$ subgraphs L for some fixed l, and recalling the definition of $e_j(H,F)$, we have

$$\sum_{L\subset H}e(L\cup F)=\sum_{j=0}^{r}\binom{k-j}{l-j}e_j(H,F).$$

Combining these and dividing by $\binom{k}{l}$ gives

$$\sum_{j=0}^{r}\frac{(l)_j}{(k)_j}e_j(H,F)=p\binom{l+\eta n}{r}+B(e^rC\eta^{-r}).$$

Writing $N_j=|\{\sigma\in(V(H)\cup V(F))^{(r)};|\sigma\cap V(H)|=j\}|$,

$$\binom{l+\eta n}{r}=\binom{k}{l}^{-1}\sum_{L\subset F}\binom{l+\eta n}{r}=\binom{k}{l}^{-1}\sum_{j=0}^{r}\binom{k-j}{l-j}N_j=\sum_{j=0}^{r}\frac{(l)_j}{(k)_j}N_j.$$

Putting $x_j=(e_j(H,F)-pN_j)/(k)_j$ we obtain

$$\sum_{j=0}^{r}(l)_jx_j=B(e^rC\eta^{-r}).$$

This equation holds for any $l\leqslant k$; selecting $r+1$ equations with $l=im$, $0\leqslant i\leqslant r$, where $rm\leqslant k$, we derive, via Lemma 2,

$$x_j=\left(\frac{2e^2}{jm}\right)^jB(2^re^rC\eta^{-r}).$$

Choosing $m=\lfloor k/r\rfloor$, then $m\geqslant k/2r$ since $k\geqslant 2r$, and we have

$$\begin{aligned}e(H)=e_r(H,F)&=pN_r+(k)_rx_r\\&=p\binom{k}{r}+B(2^{2r}e^{3r}k^r(rm)^{-r}C\eta^{-r})\\&=p\binom{k}{r}+B(e^{6r}C\eta^{-r}).\end{aligned}$$

Finally, suppose that $(1-\eta)n\leqslant k\leqslant\eta n$ (this happens only if $\eta\geqslant\frac{1}{2}$). Summing the number of edges in all subgraphs L of order $l=(1-\eta)n$ (in a similar manner to the first paragraph of the proof), and using the above we get

$$\binom{k-r}{l-r}e(H)=\binom{k}{l}\left[p\binom{l}{r}+B(e^{6r}C\eta^{-r})\right]$$

and so

$$\begin{aligned}e(H)&=p\binom{k}{r}+\frac{(k)_r}{(l)_r}B(e^{6r}C\eta^{-r})\\&=p\binom{k}{r}+B\left(\frac{e^{6r}C}{\eta^r(1-\eta)^r}\right).\quad\square\end{aligned}$$

Next we extend Lemma 3 by bounding the number of edges in a union of disjoint induced subgraphs of G. The significant feature is that this bound does not depend on the number of subgraphs, as would be the case if we were simply to apply Lemma 3 to each individual subgraph, and then sum.

Lemma 4. *Let C, n, p, η and G be as in the statement of Lemma 3, and let $s\geq 0$ be an integer. Let $H_1,\ldots,H_s$ be pairwise disjoint induced subgraphs of G, with orders $k_1,\ldots,k_s$ respectively. Then*

$$\left|\sum_{i=1}^{s}\left[e(H_i)-p\binom{k_i}{r}\right]\right|\leq Ce^{9r}\eta^{-r}(1-\eta)^{-r}/2.$$

Proof. Let $A=Ce^{6r}\eta^{-r}(1-\eta)^{-r}$. Since each summand $|e(H_i)-p\binom{k_i}{r}|$ is, by Lemma 3, bounded by A, we may suppose $s\geq e^{3r}/2$. Let H be the subgraph of G induced by $\bigcup_{i=1}^{s}V(H_i)$, and let n_j be the number of edges of H meeting exactly j of the sets $V(H_i)$, $0\leq j\leq r$. If l is some integer, $0\leq l\leq s$, and we consider all those subgraphs F of H induced by the union of some l of the $V(H_i)$, we obtain

$$\sum_F e(F)=\sum_{j=1}^{r}\binom{s-j}{l-j}n_j.$$

From Lemma 3 we have $e(F)=p\binom{|F|}{r}+B(A)$. Further, if we denote by N_j the number of r-tuples in $V(H)^{(r)}$ meeting exactly j of the sets $V(H_i)$, we get

$$\sum_F\binom{|F|}{r}=\sum_{j=1}^{r}\binom{s-j}{l-j}N_j.$$

Hence

$$\sum_{j=1}^{r}\binom{s-j}{l-j}(n_j-pN_j)=\binom{s}{l}B(A).$$

As $n_0=N_0=0$ we may extend the sum to include $j=0$, and on writing $x_j=(n_j-pN_j)/(s)_j$ we observe

$$\sum_{j=0}^{r}(l)_jx_j=B(A).$$

If $m \leqslant s/r$ is an integer, we derive $r+1$ equations by choosing $l = im$, $0 \leqslant i \leqslant r$, and then Lemma 2 yields

$$x_j = \left(\frac{2e^2}{jm}\right)^j B(2^r A).$$

Now $n_1 = \sum_{i=1}^{s} e(H_i)$ and $N_1 = \sum_{i=1}^{s} \binom{k_i}{r}$, so, choosing $m = \lfloor s/r \rfloor \geqslant s(1 - r/s)/r$, we have

$$\left|\sum_{i=1}^{s}\left[e(H_i) - p\binom{k_i}{r}\right]\right| = |n_1 - pN_1| = s\,|x_1| = \frac{2e^2 s}{m} B(2^r A)$$
$$\leqslant \frac{2e^2 r}{1 - r/s} B(2^r A) = B(e^{3r}A/2). \quad \square$$

Theorem 5. *Let n, p, α, η, ω be positive real numbers with $p, \eta < 1 \leqslant \alpha$ such that ηn is an integer with $2r \leqslant \eta n \leqslant n - 2r$. Let G be an r-uniform hypergraph $(r \geqslant 3)$ of order n in which every induced subgraph H of order ηn satisfies $|e(H) - p\binom{\eta n}{r}| \leqslant \eta n\alpha$. Then G contains a subgraph G^* of order at least*

$$(1 - e^{9r}\eta^{1-r}(1-\eta)^{-r}\omega^{-1})n$$

which is $(p, \omega\alpha)$-jumbled.

Proof. We first construct a hypergraph G_0 by repeatedly removing 'dense' subgraphs $L_1, \ldots, L_s$ such that $e(L_i) - p\binom{k_i}{r} > k_i\omega\alpha$, where $|L_i| = k_i$ and $L_j \subset G - \bigcup_{i<j} L_i$. We stop when it is no longer possible to choose another L_*, and let $G_0 = G = \bigcup_{i=1}^{s} L_i$. Let $H = \bigcup_{i=1}^{s} L_i$ and $k = |H| = \sum_{i=1}^{s} k_i$. By Lemma 4,

$$\sum_{i=1}^{s} e(L_i) \leqslant \sum_{i=1}^{s} p\binom{k_i}{r} + e^{3r}A/2,$$

where $A = e^{6r}\eta^{1-r}(1-\eta)^{-r}n\alpha$. This gives $\sum_{i=1}^{s} k_i\omega\alpha \leqslant e^{3r}A/2$ and $k \leqslant e^{3r}A/2\omega\alpha$.

Now construct G^* by removing from G_0 'sparse' subgraphs $F_1, \ldots, F_t$ such that $e(F_i) - p\binom{f_i}{r} < -f_i\omega\alpha$, where $f_i = |F_i|$. By a similar argument, we have $|G_0 - G^*| < e^{3r}A/2\omega\alpha$. Thus $|G - G^*| < e^{3r}A/\omega\alpha$, as asserted. $\square$

3. Properties of jumbled hypergraphs

We shall now explore some of the consequence of our definition of jumbled hypergraphs. In [12], properties of jumbled graphs, such as the connectivity, the number of hamilton cycles, the number of k-cliques and the contraction number, were estimated. Most the arguments, though sometimes involved, were based upon these next two simple propositions.

Proposition C ([12]). *Let G be a (p, α)-jumbled graph of order n, and let $0 < \varepsilon < 1$. Then at least $(1-\varepsilon)n$ of the vertex degrees of G lie in the range $p(n-1) \pm 10\alpha\varepsilon^{-1}$.*

Proposition D ([12]). *Let G be a (p, α)-jumbled graph of order n, and let $0 < \varepsilon < 1$. Let H be an induced subgraph of G of order k. Then at least $n - \varepsilon k$ of the vertices of G have between $pk - 21\alpha\varepsilon^{-1}$ and $pk + 21\alpha\varepsilon^{-1}$ neighbours in H.*

In this section, we shall first prove versions of these propositions for hypergraphs. For this, the following lemma is required.

Lemma 6. *Let G be a (p, α)-jumbled r-uniform hypergraph, and let S and T be any two vertex-disjoint induced subgraphs of G. Then*

$$e_i(S, T) = p\binom{s}{i}\binom{t}{r-i} + B(e^{5r}\alpha(s+t)),$$

where $s = |S|$ and $t = |T|$.

Proof. The lemma is clearly true if $\max\{s, t\} \leqslant 2r$, since by definition $\alpha \geqslant 1$. Therefore we assume otherwise, say $s \geqslant 2r$. Let L be a subgraph of S of order l, where $1 \leqslant l \leqslant s$. Then

$$e(L \cup T) = p\binom{l+t}{r} + B(\alpha(l+t)),$$

and summing over all such subgraphs L, with l fixed, gives

$$\sum_{j=0}^{r} \binom{s-j}{l-j} e_j(S, T) = \sum_L e(L \cup T),$$

whence

$$\sum_{j=0}^{r} \frac{(l)_j}{(s)_j} e_j(S, T) = p\binom{l+t}{r} + B(\alpha(l+t)).$$

Writing $Nj = \binom{s}{j}\binom{t}{r-j}$ for the number of r-tuples in $(S \cup T)^{(r)}$ with exactly j elements in S, we have also

$$\binom{l+t}{r} = \binom{s}{l}^{-1} \sum_L \binom{l+t}{r} = \binom{s}{l}^{-1} \sum_{j=0}^{r} \binom{s-j}{l-j} N_j = \sum_{j=0}^{r} \frac{(l)_j}{(s)_j} N_j.$$

Putting $x_j = (e_j(S, T) - pN_j)/(s)_j$, it follows that

$$\sum_{j=0}^{r} (l)_j x_j = B(\alpha(l+t)) = B(\alpha(s+t)).$$

If $m \leqslant s/r$ is an integer, we may obtain $r+1$ equations by setting $l = im$,

$0 \leqslant i \leqslant r$, and then Lemma 2 yields

$$x_j = \left(\frac{2e^2}{jm}\right)^j B(2^r\alpha(s+t)).$$

Thus, choosing $m\lfloor s/r \rfloor \geqslant s/2r$ since $s \geqslant 2r$,

$$\begin{aligned} e_j(S, T) &= pN_j + (s)_j x_j \\ &\leqslant pN_j + \left(\frac{2e^2 s}{jm}\right)^j B(2^r\alpha(s+t)) \\ &\leqslant pN_j + 2^{2r}e^{2r}\left(\frac{2r}{j}\right)^j B(\alpha(s+t)) \\ &\leqslant pN_j + 2^{2r}e^{2r}e^{2r/e}B(\alpha(s+t)) \\ &\leqslant p\binom{s}{j}\binom{t}{r-j} + B(e^{5r}\alpha(s+t)). \quad \square \end{aligned}$$

We are now in a position to prove an analogue of Proposition C for hypergraphs.

Lemma 7. *Let G be a (p, α)-jumbled r-uniform hypergraph ($r \geqslant 3$) of order n, with $0 < \varepsilon < 1$. Then at least $(1-\varepsilon)n$ of the vertex degrees of G lie in the range $p\binom{n-1}{r-1} \pm e^{6r}\alpha\varepsilon^{-1}$.*

Proof. Let S be a subgraph of order s, and let the sum of the degrees (in G) of the vertices of S be sd. Then

$$sd = \sum_{j=1}^{r} je_j(S, G-S),$$

and using Lemma 6 we have

$$sd = p\sum_{j=1}^{r} j\binom{s}{j}\binom{n-s}{r-j} + B\left(e^{5r}\alpha n \sum_{i=1}^{r} i\right),$$

so

$$d = p\binom{n-1}{r-1} + B(e^{6r}\alpha n/2s).$$

Thus taking S to be the $\lceil \varepsilon n/2 \rceil$ vertices of smallest degree in G, we see that the average of these degrees is at least $p\binom{n-1}{r-1} - e^{6r}\alpha\varepsilon^{-1}$. The proof is completed by taking S to be an $\lceil \varepsilon n/2 \rceil$ vertices of highest degree in G. $\square$

We also have a version of Proposition D for hypergraphs.

Lemma 8. *Let G be a (p, α)-jumbled r-uniform hypergraph ($r \geqslant 3$) of order n,* *with $0<\varepsilon<1$. Let H be an induced subgraph of G of order k. Then at least $n-\varepsilon k$ of the vertices of G have between $p\binom{k}{r-1}-e^{7r}\alpha\varepsilon^{-1}$ and $p\binom{k}{r-1}+e^{7r}\alpha\varepsilon^{-1}$ neighbours in H.*

Proof. By Lemma 7 applied to the (p, α)-jumbled hypergraph H, at most $\varepsilon k/3$ vertices of H have degrees in H outside the specified range. Let S be a set of s vertices of $G-H$, and let d be the average degree in H of the vertices in S. Then, by Lemma 6,

$$sd = e_1(S, V(H)) = ps\binom{k}{r-1} + B(e^{5r}\alpha(s+k)).$$

Hence $d=p\binom{k}{r-1}+B(e^{5r}\alpha(1+k/s))$. Choosing S to be the $\lceil \varepsilon k/3 \rceil$ vertices of $G-H$ of highest degree in H, we see that all but $\varepsilon k/3$ vertices of $G-H$ have degree at most $p\binom{k}{r-1}+e^{7r}\alpha\varepsilon^{-1}$ in H. A similar argument applied to the vertices of $G-H$ with low degree in H completes the proof. □

Several of the graph properties studied in [12] have hypergraph analogues. For instance it is easily seen, by a crude estimate, that the clique and independence numbers of a (p, α)-jumbled hypergraph are at most $\alpha^{1/(r-1)}$, whence the chromatic number is at least $n\alpha^{-1/(r-1)}$. Of more interest is a lower bound on the clique number. For $(\frac{1}{2}, n^{\frac{1}{2}})$-jumbled graphs, the following proposition from [12], with $F=K_k$, showed that for k up to about $(\log_2 n)/2$, the number of k-cliques is approximately that found in a random graph, and so in particular the clique number is at least $(\log_2 n)/2$.

Proposition E ([12]). *Let F be a graph of order $r\geqslant 3$ with m edges, and let z be the order of its automorphism group. Let G be a (p, α)-jumbled graph of order n, where $p\leqslant\frac{1}{2}$. Suppose ε satisfies $0<\varepsilon<1$ and $\varepsilon^2p^rn\geqslant 42\alpha r^2$. Then the number of induced subgraphs of G which are isomorphic to F lies between $(1-\varepsilon)^rp^mq^{\binom{r}{2}-m}z^{-1}n^r$ and $(1+\varepsilon)^rp^mq^{\binom{r}{2}-m}z^{-1}n^r$, where $q=1-p$.*

It would be desirable to have a result for hypergraphs in the spirit of Proposition E. A specific reason for doing so, apart from its yielding a lower bound for the clique number, would be to solve this next problem of Erdös and Sós, posed in [7]:

Problem. Let H be an r-uniform hypergraph and $f(n; H)$ be the smallest integer for which every r-uniform hypergraph of n vertices and more than $f(n; H)$ edges contains a subgraph isomorphic to H. An *extremal graph belonging to H* is a hypergraph G with $e(G)=f(|G|; H)$ which does not contain a subgraph isomorphic to H. We define a sequence of hypergraphs G_i $(i=1,2,\ldots)$ to be *uniformly distributed* if $|G_i|=i$, and for every $\eta>0$ there is a $c(\eta)$, so that for every $i > i_0(\eta)$ every induced subgraph of G_i with $m > \eta i$ vertices has $(c(\eta) + o(1))\binom{m}{r}$ edges. Is it true that there is no sequence of extremal graphs belonging to H which is uniformly distributed? (In particular, is it true for the case $H = K_4^3$, the complete 3-uniform hypergraph of order 4?)

The proof of Proposition E (with $F = K_k$) goes roughly as follows. Select a vertex x_1 and let H_1 be the subgraph spanned by its neighbours. For most choices of x_1, $|H_1| \approx pn$. Select a vertex x_2 of H_1 and let H_2 be the subgraph spanned by the neighbours of x_2 in H_1. Again, for most choices of x_2, $|H_2| \approx p^2 n$, and so on. In this way ordered k-cliques $\{x_1, \ldots, x_k\}$ are counted. To be able to count cliques in a jumbled hypergraph, we would need something to the effect that for each vertex, the $(r-1)$-uniform hypergraph induced by its neighbours was jumbled, and that this $(r-1)$-uniform hypergraph was in some sense 'independent' of the original hypergraph. Such properties will not hold for the general jumbled hypergraph, though even if they do, the analogue of Proposition E may still fail; here is a class of examples.

A *division* of the set $X = \{1, \ldots, n\}$ will be a collection $\mathcal{F}$ of functions f_ξ, $\xi \in X^{(r-2)}$, such that $f_\xi : (X - \xi) \to \{-1, 1\}$. The set of all divisions is given the uniform probability distribution, so each division has probability 2^{-m}, $m = (n - r + 2)\binom{n}{r-2}$. Let p and δ be real numbers, $0 < p$, $\delta < 1$. For each $\xi \in X^{(r-2)}$ and $\{x, y\} \in (X - \xi)^{(2)}$ we define the random variable $\varepsilon(\{x, y\}; \xi) = -f_\xi(x) f_\xi(y)\delta$. Thus $\varepsilon(\{x, y\}; \xi)$ equals δ if $f_\xi(x) \neq f_\xi(y)$ and equals $-\delta$ otherwise. A given division $\mathcal{F}$ induces a probability distribution on the set of r-uniform hypergraphs with vertex set X as follows: the edges appear independently, and for $\sigma \in X^{(r)}$,

$$\Pr(\sigma \text{ is an edge, given } \mathcal{F}) = 1 - \prod_{\xi \in \sigma^{(r-2)}} (1-p)^{1/\binom{r}{2}}(1 + \varepsilon(\sigma - \xi; \xi)) = g(\sigma, \mathcal{F}).$$

Observe if $\delta = 0$ this probability equals p. We can think of σ as being an edge as a result of at least one success among a set of Bernoulli trials, one for each $\xi \in \sigma^{(r-2)}$, each with probability of failure $(1-p)^{1/\binom{r}{2}}(1 + \varepsilon(\sigma - \xi; \xi))$. (We will assume δ is sufficiently small that, say, $(1-p)^{1/\binom{r}{2}}(1 + \delta) \leq 1$). We define the space $\mathcal{H}_r(n, p, \delta)$, which is the set of r-uniform hypergraphs with vertex set X, wherein, for a given set $A \subset X^{(r)}$,

$$\Pr(A \subset E(G)) = \sum_{\mathcal{F}} \Pr(\mathcal{F}) \prod_{\sigma \in A} g(\sigma, \mathcal{F}).$$

Hence the probability of generating a given hypergraph G is the expected value of the probability of G given $\mathcal{F}$. More wieldy expressions for $\Pr(A \subset E(G))$ are given by the next lemma; prior to stating it, we require some more notation.

For $A \subset X^{(r)}$ and $\xi \in X^{(r-2)}$, the graph A_ξ has vertex set X and edge set $\{\sigma - \xi; \sigma \supset \xi \text{ and } \sigma \in A\}$. Further, let $T_\xi(A)$ denote the set of eulerian subgraphs of A_ξ (those in which all the vertex degrees are even). Finally, let $T(A) = \prod_\xi T_\xi(A)$, and for $t = \prod_\xi t_\xi \in T(A)$ let $\#t = |\bigcup_\xi \{\sigma \in X^{(r)}; \sigma - \xi \in E(t_\xi)\}|$. Thus

$\#t$ is the number of σ in A needed to construct all the eulerian graphs $t_\xi \subset A_\xi$ which form the components of t.

Lemma 9. *Let $A \subset X^{(r)}$. In the probability space $\mathcal{H}_r(n, p, \delta)$,*

$$\Pr(A \subset E(G)) = p^{|A|} \sum_{t \in T(A)} \left(\frac{p-1}{p}\right)^{\#t} \prod_\xi (-\delta)^{e(t_\xi)}$$

$$= \sum_{B \subset A} (p-1)^{|B|} \prod_\xi \sum_{C \in T_\xi(B)} (-\delta)^{e(C)}.$$

Proof. We have

$$\Pr(A \subset E(G)) = \sum_{\mathcal{F}} \Pr(\mathcal{F}) \prod_{\sigma \in A} \left[1 - (1-p) \prod_{\xi \in X^{(r-2)}} (1 + \varepsilon(\sigma - \xi; \xi))\right]$$

$$= \boldsymbol{E}\left(\prod_{\sigma \in A} \left[p + (p-1) \sum_{0 \neq Y \subset \sigma^{(r-2)}} \prod_{\xi \in Y} \varepsilon(\sigma - \xi; \xi)\right]\right),$$

where $\boldsymbol{E}$ denotes expectation. Define $Q(A)$ to be $\{(\sigma, \xi); \sigma \in A, \xi \in \sigma^{(r-2)}\}$, and for each $R \subset Q(A)$ define R_ξ to be $R \cap (X^{(r)} \times \{\xi\})$ for each $\xi \in X^{(r-2)}$, and $\#R$ to be $|\{\sigma; (\sigma, \xi) \in R \text{ for some } \sigma\}|$. Then

$$\Pr(A \subset E(G)) = \boldsymbol{E}\left(\sum_{R \subset Q(A)} p^{|A| - \#R}(p-1)^{\#R} \prod_\xi \prod_{\sigma \in R_\xi} \varepsilon(\sigma - \xi; \xi)\right)$$

$$= p^{|A|} \sum_{R \subset Q(A)} \left(\frac{p-1}{p}\right)^{\#R} \prod_\xi \boldsymbol{E}\left(\prod_{\sigma \in R_\xi} \varepsilon(\sigma - \xi; \xi)\right),$$

because the values of $\varepsilon(*; \xi)$ and $\varepsilon(*; \xi')$ are independent if $\xi \neq \xi'$. Now R_ξ corresponds in an obvious way to a subgraph B_ξ of A_ξ. Given a vertex v of this graph, we may partition the divisions into pairs $(\mathcal{F}, \mathcal{F}')$, such that $f_\xi(v) = -f'_\xi(v)$ and $\mathcal{F}$ and $\mathcal{F}'$ otherwise agree. If the degree of v is odd, the value of $\prod_{\sigma \in R_\xi} \varepsilon(\sigma - \xi; \xi)$, given $\mathcal{F}$, will be minus one times its value under $\mathcal{F}'$. Hence the final expectation will vanish unless B_ξ is eulerian. On the other hand, if B_ξ is eulerian, then for every f_ξ,

$$\prod_{\sigma \in R_\xi} \varepsilon(\sigma - \xi; \xi) = \prod_{uv \in E(B_\xi)} -f_\xi(u) f_\xi(v) \delta = (-\delta)^{e(B_\xi)},$$

since the number of edges between $f_\xi^{-1}(-1)$ and $f_\xi^{-1}(1)$ is even. Now, because every R_ξ corresponds to an eulerian subgraph, we see that R_ξ corresponds to $t \in T(A)$ and $\#R = \#t$. Thus

$$\Pr(A \subset E(G)) = p^{|A|} \sum_{t \in T(A)} \left(\frac{p-1}{p}\right)^{\#t} \prod_\xi (-\delta)^{e(t_\xi)}.$$

Finally, on writing $\sigma(\boldsymbol{t})=\bigcup_{\xi}\{\sigma\in X^{(r)};\ \sigma-\xi\in E(t_{\xi})\}$, so that $\#\boldsymbol{t}=|\sigma(\boldsymbol{t})|$, we see

$$\begin{aligned}
\Pr(A\subset E(G)) &= p^{|A|}\sum_{Y\subset A}\left(\frac{p-1}{p}\right)^{|Y|}\sum_{\substack{\boldsymbol{t}\\ \sigma(\boldsymbol{t})=Y}}\prod_{\xi}(-\delta)^{e(t_{\xi})}\\
&= p^{|A|}\sum_{Y\subset A}\left(\frac{p-1}{p}\right)^{|Y|}\sum_{\substack{\boldsymbol{t}\\ \sigma(\boldsymbol{t})\subset Y}}\left(\sum_{\substack{B\\ \sigma(\boldsymbol{t})\subset B\subset Y}}(-1)^{|B|-|Y|}\right)\prod_{\xi}(-\delta)^{e(t_{\xi})}\\
&= p^{|A|}\sum_{Y\subset A}\left(\frac{1-p}{p}\right)^{|Y|}\sum_{B\subset Y}(-1)^{|B|}\sum_{\substack{\boldsymbol{t}\\ \sigma(\boldsymbol{t})\subset B}}\prod_{\xi}(-\delta)^{e(t_{\xi})}\\
&= p^{|A|}\sum_{B\subset A}(-1)^{|B|}\left(\sum_{B\subset Y\subset A}\left(\frac{1-p}{p}\right)^{|Y|}\right)\prod_{\xi}\sum_{C\in T_{\xi}(B)}(-\delta)^{e(C)}\\
&= \sum_{B\subset A}(p-1)^{|B|}\prod_{\xi}\sum_{C\in T_{\xi}(B)}(-\delta)^{e(C)}.\quad\square
\end{aligned}$$

We are now in a position to show that the graphs in $\mathcal{H}_r(n,p,\delta)$ are most surely jumbled; in fact a considerably stronger statement is true.

Theorem 10. *Almost every hypergraph in $\mathcal{H}_r(n,p,\delta)$ has the property that, for each l, $0\leqslant l\leqslant r-3$, and for $v\in X^{(l)}$, the $(r-l)$-uniform hypergraph induced on $X-v$ by the edges containing v is $(p,O(n^{r-l-\frac{5}{4}}\log n))$-jumbled.*

Proof. It is sufficient to prove the statement for $l=r-3$, for then, if $v\in X^{(l)}$ and H is a subgraph of the induced subgraph on $X-v$, where $|H|=k$, consider the 3-uniform hypergraph H_ρ induced on $V(H)-\rho$ by $v\cup\rho$, where $\rho\in V(H)^{(r-3-l)}$. We have

$$\begin{aligned}
e(H) &= \binom{r-l}{r-l-3}^{-1}\sum_{\rho}e(H_\rho)\\
&= \binom{r-l}{r-l-3}^{-1}\binom{k}{r-l-3}\left[p\binom{k-r+3-l}{3}+O(kn^{\frac{7}{4}}\log n)\right]\\
&= p\binom{k}{r-l}+O(kn^{r-l-\frac{5}{4}}\log n).
\end{aligned}$$

So, let $G\in\mathcal{H}_r(n,p,\delta)$, let $v\in X^{(r-3)}$, and let G_v be the 3-uniform hypergraph induced on $X-v$ by v. Let $x,y\in X-v$ and $Z\subset X-v-\{x,y\}$. Setting $A=\{v\cup\{x,y,z\};\ z\in Z\}$, we see that for every $\xi\in X^{(r-2)}$, A_ξ is empty or is a star. Since A_ξ contains no non-empty eulerian subgraphs, Lemma 9 implies

$$\Pr(\{x,y,z\}\in E(G_v);\ z\in Z)=\Pr(A\subset E(G))=p^{|A|}=p^{|Z|},$$

so the occurrence of edges of G_v containing $\{x,y\}$ follows a binomial distribution. By standard estimates, the number of edges containing $\{x,y\}$ lies in the range

$pn \pm n^{\frac{1}{2}} \log n$ with probability $1 + O(n^{-\log n})$, so with probability $1 + O(n^{2-\log n})$ every 2-tuple of $V(G_v)$ is contained in $pn \pm n^{\frac{1}{2}} \log n$ edges of G_v.

We now estimate the number of common neighbours of x and y. As the graph K_{n-r-1} has edge chromatic number at most $n - r - 1$, the set $(X - v - \{x, y\})^{(2)}$ can be partitioned into sets $M_1, \ldots, M_{n-r-1}$ such that for each i and $\lambda, \mu \in M_i$, $\lambda \cap \mu = \emptyset$ holds. Moreover $\lfloor (n-r-1)/2 \rfloor \le |M_i| \le \lceil (n-r-1)/2 \rceil$. Let $W \subseteq M_i$ and let $A = \{v \cup \{x\} \cup \lambda, v \cup \{y\} \cup \lambda; \lambda \in W\}$. For any $\xi \in X^{(r-2)}$, A_ξ is empty, a path, or a set of independent edges, because the $\lambda \in W$ are disjoint. Hence, once again,

$$\Pr(\{x\} \cup \lambda, \{y\} \cup \lambda \in E(G_v); \lambda \in W) = \Pr(A \subset E(G_v)) = p^{|A|} = p^{2|W|},$$

so the number of $\lambda \in W$ with $\{x\} \cup \lambda$ and $\{y\} \cup \lambda$ in $E(G)$ follows a binomial distribution with probability p^2. Thus the number of λ in M_i with this property is at most $p^2 n/2 + n^{\frac{1}{2}} \log n$ with probability $1 + O(n^{-\log n})$, and summing over all M_i we see the number of common neighbours of x and y is at most $p^2\binom{n-2}{2} + n^{\frac{3}{2}} \log n$ with probability $1 + O(n^{1-\log n})$. The same holds for all pairs $\{x, y\} \in (X - v)^{(2)}$ with probability $1 + O(n^{3-\log n})$.

Applying Theorem 1 to G_v, we see that G_v is $(p, O(n^{\frac{7}{4}} \log n))$-jumbled with probability $1 + O(n^{3-\log n})$, and this will hold for every $v \in X^{(r-3)}$ with probability $1 + O(n^{r-\log n}) = 1 + o(1)$, as claimed. □

Theorem 10 cannot be extended to $l = r - 2$. For let $v \in X^{(r-2)}$, and consider G_v and $\{x, y\} \in (X - v)^{(2)}$ as before. If $Z \subset X - v - \{x, y\}$ and $A = \{v \cup \{x\} \cup z, v \cup \{y\} \cup z; z \in Z\}$, then A_v is a complete bipartite graph $K_{2,|Z|}$. From this it follows that

$$\begin{aligned}\Pr(A \subset E(G)) &= p^{|A|} \sum_{\substack{j=0 \\ j \text{ even}}}^{|A|} \binom{|A|}{j} \left(\frac{p-1}{p}\right)^{2j} (-\delta)^{2j} \\ &= \tfrac{1}{2}\{(p^2 + (1-p)^2\delta^2)^{|A|} + (p^2 - (1-p)^2\delta^2)^{|A|}\}.\end{aligned}$$

Hence the distribution of common neighbours of x and y is bimodal, being the average of two binomial distributions with probabilities $p^2 + (1-p)^2\delta^2$ and $p^2 - (1-p)^2\delta^2$. Certainly the number of common neighbours will almost surely not lie close to $p^2 n$, so G_v will not be $(p, o(n))$-jumbled. To some extent, this may explain why the proportion of $(r+1)$-cliques in $G \in \mathcal{H}_r(n, p, \delta)$ is not that found in a random hypergraph, namely p^{r+1}, as we proceed to demonstrate.

Theorem 11. *Let $K \subset X$, $|K| = r + 1$. Then for $G \in \mathcal{H}_r(n, p, \delta)$,*

$$\begin{aligned}\Pr(K^{(r)} \subset G) &= \sum_{j=0}^{r+1} \binom{r+1}{j} (p-1)^j (1-\delta^3)^{\binom{j}{3}} \\ &= p^{r+1} + p^{r-2}(1-p)^3\delta^3 + O(\delta^6),\end{aligned}$$

the last term signifying p constant and $\delta \to 0$. In particular there are values of δ for which $\Pr(K^{(r)} \subset G) \neq p^{r+1}$.

Proof. Let $B \subset K^{(r)}$, $|B| = j$. Then $B = \{K - \{y\}; y \in Y\}$ for some set Y with $|Y| = j$. Let $\xi \in X^{(r-2)}$. The graph B_ξ is empty unless $\xi \in K^{(r-2)}$, in which case $\rho = K - \xi \in K^{(3)}$ and B_ξ has one edge for each element of $Y \cap \rho$. So B_ξ contains no non-trivial eulerian subgraphs unless $\rho \subset Y$, when B_ξ is a triangle. There are $\binom{j}{3}$ such ρ, and hence ξ, for which this holds. By Lemma 9,

$$\Pr(K^{(r)} \subset G) = \sum_{B \subset K^{(r)}} (p-1)^{|B|} \prod_{\rho \in Y^{(3)}} (1 - \delta^3)$$
$$= \sum_{j=0}^{r+1} \binom{r+1}{j} (p-1)^j (1-\delta^3)^{\binom{j}{3}}. \quad \square$$

Although we cannot prove that the proportion of $(r+1)$-cliques in a $(p, o(n^{r-1}))$-jumbled hypergraph is around p^{r+1}, it may yet be possible to establish that the number of $(r+1)$-cliques in non-zero. This would be enough to answer the above-mentioned question of Erdős and Sós.

Appendix

Here we establish the identities employed in the proof of Theorem 1.

Identity 1.

$$-\frac{1}{(k-r+1)} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 S(j)\, |X_j|^{-1} = \binom{k}{r-1} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1}$$
$$= \binom{n-r+2}{r-1} \binom{n-k}{r-1}^{-1}.$$

Proof. We start by demonstrating that the first two expressions are equivalent. Taking the left-hand side, we have

$$\frac{k}{(k-r+1)} \binom{n-1}{r-1} \binom{n-1}{k-1}^{-1} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} \sum_{i=1}^{j} (-1)^{i-1} \binom{n-r+1}{k-r+i}$$
$$= \binom{k}{r-1} \binom{n-r}{k-r}^{-1} \sum_{j=1}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} \sum_{i=1}^{j} (-1)^{i-1} \binom{n-r+1}{k-r+i}$$
$$= \binom{k}{r-1} \binom{n-r}{k-r}^{-1} \sum_{j=1}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} \sum_{i=1}^{j} (-1)^{i-1} \left[\binom{n-r}{k-r+i-1} + \binom{n-r}{k-r+i} \right]$$
$$= \binom{k}{r-1} \binom{n-r}{k-r}^{-1} \sum_{j=1}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} \left[\binom{n-r}{k-r} + (-1)^{j-1} \binom{n-r}{k-r+j} \right]$$

$$= \binom{k}{r-1} \sum_{j=1}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} + \frac{1}{k-r+1} \sum_{j=1}^{r-1} (-1)^{j-1} \binom{r-1}{j} (k-r+j+1)$$
$$= \binom{k}{r-1} \sum_{j=1}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} + 1$$
$$= \binom{k}{r-1} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1} \quad \text{as required.}$$

To prove that the second two expressions are equivalent, we observe, by expanding binomial coefficients and rearranging,

$$\binom{k}{r-1}\binom{n-k}{r-1} \sum_{j=0}^{r-1} \binom{r-1}{j}^2 |X_j|^{-1}$$
$$= \sum_{j=0}^{r-1} \binom{k-r+j+1}{j} \binom{n-k-j}{n-k-r+1} = \binom{n-r+2}{r-1},$$

the last step following by counting the number of geodesics in $\mathbb{Z}^2$ from $(0,0)$ to $(r-1, n-2r+3)$ passing through $(j, k-r+2)$. $\square$

Identity 2.

$$\frac{1}{k} \sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} S(j) = \binom{n-1}{r-1}.$$

Proof. Taking the left-hand side of the expression, we have

$$\binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{j=1}^{r-1} \binom{r-1}{j} \sum_{i=1}^{j} (-1)^{j-i} \binom{n-r+1}{k-r+i}$$
$$= \binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i=1}^{r-1} \binom{n-r+1}{k-r+i} \sum_{j=i}^{r-1} \binom{r-1}{j} (-1)^{j-i}$$
$$= \binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i=1}^{r-1} \binom{n-r+1}{k-r+i} \sum_{j=i}^{r-1} \left[\binom{r-2}{j-1} + \binom{r-2}{j}\right] (-1)^{j-i}$$
$$= \binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i=1}^{r-1} \binom{n-r+1}{k-r+i} \binom{r-2}{i-1}$$
$$= \binom{n-1}{r-1}\binom{n-1}{k-1}^{-1} \sum_{i=0}^{r-2} \binom{n-r+1}{n-k-i} \binom{r-2}{i}$$
$$= \binom{n-1}{r-1}. \quad \square$$

Identity 3.

$$\sum_{j=0}^{r-1} \binom{r-1}{j}^2 S(j)^2 |X_j|^{-1} - (k-r+1)^2 \binom{k}{r-1} \binom{n-r+2}{r-1} \binom{n-k}{r-1}^{-1}$$
$$= k(k-1)\binom{n-2}{r-1} + k\binom{n-1}{r-1}.$$

Proof. Taking the left-hand side of the expression, expanding the binomial coefficients in the first term, and multiplying its numerator and denominator by $(k-r+1)!\,(n-k-r+1)!$, we obtain

$$\begin{aligned}
&\frac{(k-r+1)!\,(n-k-r+1)!\,k!\,(n-k)}{(n-r)!^2}\sum_{j=0}^{r-1}\binom{k-r+j+1}{j}\binom{n-k-j}{n-k-r+1}\\
&\quad\times\left(\sum_{i=1}^{j}(-1)^i\binom{n-r+1}{k-r+i}\right)^2-(k-r+1)^2\binom{k}{r-1}\binom{n-r+2}{r-1}\binom{n-k}{r-1}^{-1}\\
&=\left\{\sum_{j=0}^{r-1}\binom{k-r+j+1}{j}\binom{n-k-j}{n-k-r+1}\right.\\
&\quad\times\left(\sum_{i=1}^{j}(-1)^i\left[\binom{n-r}{k-r+i-1}+\binom{n-r}{k-r+i}\right]\right)^2\\
&\quad\left.-\binom{n-r+2}{r-1}\binom{n-r}{n-k}^2\right\}\binom{k}{r-1}(k-r+1)^2\binom{n-k}{r-1}^{-1}\binom{n-r}{n-k}^{-2}\\
&=\left\{\sum_{j=0}^{r-1}\binom{k-r+j+1}{j}\binom{n-k-j}{n-k-r+1}\left((-1)^j\binom{n-r}{k-r+j}-\binom{n-r}{k-r}\right)^2\right.\\
&\quad\left.-\binom{n-r+2}{r-1}\binom{n-r}{n-k}^2\right\}\binom{k}{r-1}(k-r+1)^2\binom{n-k}{r-1}^{-1}\binom{n-r}{n-k}^{-2}.
\end{aligned}$$

Note that we have the identity $\sum_{j=0}^{r-1}\binom{k-r+j+1}{j}\binom{n-k-j}{n-k-r+1}=\binom{n-r+2}{r-1}$ from the proof of identity 2, so making this substitution, our expression becomes

$$\begin{aligned}
&\left\{\sum_{j=0}^{r-1}\binom{k-r+j+1}{j}\binom{n-k-j}{n-k-r+1}\binom{n-r}{n-k-j}\left(\binom{n-r}{n-k-j}-2(-1)^j\binom{n-r}{n-k}\right)\right\}\\
&\quad\times\binom{k}{r-1}(k-r+1)^2\binom{n-k}{r-1}^{-1}\binom{n-r}{n-k}^{-2}\\
&=k!\,(n-k)!\sum_{j=0}^{r-1}\frac{(k-r+j+1)}{j!\,(r-j-1)!}\left(\frac{1}{(n-k-j)!\,(k-j+r)!}-\frac{2(-1)^j}{(n-k)!\,(k-r)!}\right)\\
&=k!\,(n-k)!\sum_{j=0}^{r-1}\frac{(k-r+j+1)}{j!\,(r-j-1)!\,(n-k-j)!\,(k-j+r)!}\\
&\quad-2k\binom{k-1}{r-1}\sum_{j=0}^{r-1}(-1)^j\binom{r-1}{j}(k-r+j+1)\\
&=k!\,(n-k)!\sum_{j=0}^{r-1}\frac{(k-r+j+1)}{j!\,(r-j-1)!\,(n-k-j)!\,(k-j+r)!}\\
&=\frac{k!\,(n-k)!}{(r-1)!\,(n-r)!}\sum_{j=0}^{r-1}\binom{r-1}{j}\binom{n-r}{n-k-j}(k-r+j+1)\\
&=\frac{k!\,(n-k)!}{(r-1)!\,(n-r)!}\left[\sum_{j=0}^{r-1}\binom{r-1}{j}\binom{n-r-1}{n-k-j}(n-r)+\sum_{j=0}^{r-1}\binom{r-1}{j}\binom{n-r}{n-k-j}\right]
\end{aligned}$$

$$= \frac{k!\,(n-k)!}{(r-1)!\,(n-r)!}\left[(n-r)\binom{n-2}{n-k} + \binom{n-1}{n-k}\right]$$

$$= k(k-1)\binom{n-2}{r-1} + k\binom{n-1}{r-1}. \quad \square$$

References

[1] B. Bollobás, Random Graphs (Academic Press, London, 1985).
[2] B. Bollobás and A. Thomason, Parallel sorting, Discrete App. Math. 6 (1983) 1–11.
[3] B. Bollobás and W.F. de la Vega, The diameter of random regular graphs, Combinatorica 2 (1982) 125–134.
[4] F.R.K. Chung, On concentrators, superconcentrators, generalisers and non-blocking networks, Bell Syst. Tech. J. 58 (1978) 1765–1777.
[5] F.R.K. Chung, R.L. Graham and R.M. Wilson, Quasi-random graphs (preprint).
[6] P. Erdös, Some remarks on the theory of graphs, Bull. Amer. Math. Soc. 53 (1947) 292–294.
[7] P. Erdös and V.T. Sós, On Ramsey-Turán type theorems for hypergraphs, Combinatorica 2 (1982) 289–295.
[8] P. Erdös and J. Spencer, Imbalances in k-colorations, Networks 1 (1972) 379–385.
[9] J. Haviland, Ph.D. thesis, University of Cambridge (in preparation).
[10] P. Kövári, V.T. Sós and P. Turán, On a problem of K. Zarankiewicz, Colloq. Math. 3 (1954) 50–57.
[11] A. Thomason, An extremal function for contractions of graphs, Math. Proc. Cambridge Phil. Soc. 95 (1984) 261–265.
[12] A. Thomason, Pseudo-random graphs, in Proceedings of Random Graphs, Poznań 1985, M. Karonski, ed., Annals of Discrete Math. 33 (1987) 307–331.
[13] A. Thomason, Random graphs, strongly-regular graphs and pseudo-random graphs, in Surveys in Combinatorics 1987, C. Whitehead, ed., London Math. Soc. Lecture Note Series 123 (1987) 173–195.